PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 1/00, 5/10, 15/54, C12Q 1/48, 1/68

A1

(11) International Publication Number: **WO 98/35016**

(43) International Publication Date: 13 August 1998 (13.08.98)

(21) International Application Number: PCT/US98/02494

(22) International Filing Date: 9 February 1998 (09.02.98)

(30) Priority Data:
60/037,561 11 February 1997 (11.02.97) US

(71) Applicant *(for all designated States except US)*: MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: DUONG, Le, T. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). RODAN, Gideon, A. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PROTEIN TYROSINE KINASE 2 (PYK2), NUCLEIC ACIDS, AND ASSAY

(57) Abstract

This invention is directed to nucleic acids encoding protein tyrosine kinase 2 (PYK2), to murine PYK2, to methods of making this protein using the nucleic acids, and to assays for inhibitors of PYK2.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TITLE OF THE INVENTION
PROTEIN TYROSINE KINASE 2 (PYK2), NUCLEIC ACIDS, AND ASSAY

BRIEF DESCRIPTION OF THE INVENTION

This invention is directed to nucleic acids encoding protein tyrosine kinase 2 (PYK2), to murine PYK2, to methods of making this protein using the nucleic acids, and to assays for inhibitors of PYK2.

BACKGROUND OF THE INVENTION

Protein tyrosine kinase 2 (PYK2), also known as Cell Adhesion Kinase β (CAKβ) and Related Adhesion Focal Tyrosine Kinase (RAFTK) is a recently described member of the focal adhesion kinase family (Avraham *et al.*, 1995 *J. Biol. Chem.* 270:27742-27751; Lev *et al.*, 1995 *Nature*. 376:737-745; and Sasaki, *et al.*, 1995 *J. Biol. Chem.* 270:21206-21219.). PYK2 was first cloned from human brain as a Grb-2 binding protein, and has also been cloned from rat and human brain libraries. There have been conflicting reports as to its cellular expression. In one study, abundant PYK2 transcripts were found in brain and lower levels were detected in the kidney. In another report, PYK2 expression was also found to be most abundant in rat brain, but its transcripts could also be detected in kidney, spleen, lung, intestine and epididymis. PYK2 transcripts were also detected in rat fibroblast 3Y1 and WFB cell lines, as well as in the human T cell leukemia Jurkat line. When cloned from the human megakaryocytic CMK cell line and from mouse brain, it was found to have wider tissue distribution beyond brain, notably spleen, lung, thymus and peripheral blood leukocytes. In addition, expression of PYK2 was reported in human CD34+ marrow cells, megakaryocytes and platelets.

DETAILED DESCRIPTION OF THE INVENTION

One aspect of this invention is are nucleic acids, substantially free from associated nucleic acids, which encode murine protein tryosine kinase 2 (PYK2). In one embodiment, the nucleic acid which encodes PYK2 is a DNA.

Another aspect of this invention is murine PYK2 cDNA, Murine PYK2 DNA is set forth in Figure 1 (SEQ ID NO:5).

Yet another aspect of this invention is murine PYK2 which is free from associated murine proteins. One murine PYK2 is set forth in Figure 1 (SEQ ID NO:6).

Another aspect of this invention is a method of making PYK2 by introducing nucleic acids into a cell, the nucleic acids comprising nucleic acids which encode PYK2, under conditions which transcription and translation of PYK2 occur. It is preferred that the nucleic acids be present in a vector, such as a plasmid or baculovirus vector. It is also preferred that the nucleic acids be under the control of transcriptional control elements, such as promoters and optionally enhancers. Such control elements are well known in the art.

Host cells which express PYK2 are also part of this invention. Preferred host cells include mammalian cells, insect cells, yeast and bacterial cells such as *E. coli*. Cell lines which permanently (rather than transiently) express murine PYK2 are also another aspect of this invention.

The recombinant PYK2, which is made using the cloning process of this invention may be used in assays in order to further characterize the biological function of PYK2 and to identify compounds such as agonists and antagonists which modulate its activity. A further aspect of this invention is an assay for the identification of compounds which modulate the activitiy of PYK2, and particularly inhibitors of PYK2 activity. This assay comprises the steps of: contacting recombinant PYK2 with a tyrosine substrate in the presence of radiolabeled ATP and a putative activity-modifying compound, and measuring the amount of radiolabeled tyrosine which is formed; and optionally comparing the amount of radiolabeled tyrosine formed in the

presence of the putative activity-modifying compound with that formed in the absence of the putative activity-modifying compound.

Integrins are the major family of cell surface receptors that mediate adhesive interactions, either to adjacent cells or to the extracellular matrix. Integrin signalling is mediated through the focal adhesion kinase (FAK) family of proteins. PYK2 is a member of the FAK family, and is involved in integrin-mediated signal transduction pathways in megakaryocytes, brain tissue and hematopoietic cells. Modulators of PYK2 would therefore be potential therapeutic agents for modulating platelet levels.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is the cDNA sequence of mouse PYK2 and the deduced protein sequence. Intron sequences are in lower case letters. The exon sequence is capitalized. The boxed sequence of the deduced protein indicates the kinase domain. The circled prolines of the deduced protein indicate the proline rich domain.

As used throughout the specification and claims, the following definitions apply:

"PYK2" means protein tyrosine kinase 2, allelic variations of protein tyrosine kinase 2, and mutations or fragments thereof which retain at least about 85%, and preferably at least about 90% of the biological activity of native PYK2.

"Native PYK2" means the protein tyrosine kinase which is naturally occuring in an organism.

"Substantially free from associated nucleic acids" means that in a sample, there is less than about 5% (by weight) nucleic acids present which are other than nucleic acids encoding PYK2.

"Substantially free from associated murine proteins" means that in a sample, there is less than about 5% (by weight) protein which is other than murine PYK2.

"FAK" means focal adhesion kinase.

"Heterologous" PYK2 nucleic acid means that the nucleic acid was introduced to the cell, without regard as to whether the nucleic acid is from the same species as the cell; alternatively it refers to nucleic

acids encoding PYK2 in a cell whose ancestor had PYK2 introduced into the cell.

"Heterologous" PYK2 protein means that the PYK2 was encoded by a heterologous nucleic acid.

FAK proteins, which are involved in cell adhesion processes, were not detected in a number of macrophage cell lines and it was therefore hypothesized that another cell adhesion-dependent kinase, homologous to FAK, may assume its function in these cells. PYK2 was recently identified as another member of the FAK family and its expression was detected in spleen, thymus, lung and peripheral blood leukocytes (Avraham, *et al.*, 1995 *supra*; Sasaki, *et al.*, 1995 *supra*). To evaluate PYK2 as a possible adhesion-dependent kinase in macrophages, specific probes were generated for PYK2 and FAK which were used to examine the expression of PYK2 and FAK in mouse tissues. As previously reported, for other species, PYK2 is highly expressed in brain and spleen, and at lower levels in kidney, lung and liver and has a more restricted tissue distribution than FAK.

Using a PYK2 probe, the full length cDNA was cloned from a mouse spleen cDNA library. The deduced amino acid sequence of the full length clone was found to be identical to the recently published amino acid sequence of the mouse RAFTK (Avraham, *et al.*, 1995, *supra*).

In addition, the full length FAK was cloned from a mouse osteoblastic MB1.8 cell line (Wesolowski, *et al.*, 1995, *Exp. Cell Res.*, 219: 679-686.).

PYK2 and FAK cDNAs were subsequently transfected into human embryonic kidney (HEK) 293 cells. Cell lines which permanently express either PYK2 or FAK were established and the expression levels of the exogeneously expressed mouse kinases were assessed by northern analysis.

The following Examples are presented to better illustrate the invention.

EXAMPLE 1

Cell Culture

Macrophages were induced by thioglycolate injection into the peritoneal cavities of adult BALB/c mice. After 4 days, cells were collected, washed and cultured in RPMI 1640 medium containing 10% FBS. After 3h at 37°C, the cultures were washed extensively to remove non-adherent cells and cultured overnight before samples were prepared for immunoprecipitation. Bone marrow derived macrophages were prepared as described by Li and Chen, 1995 *J. Leuk. Biol.* 57:484-490, which is hereby incorporated by reference. Non adherent cells were cultured in RPMI completed medium in the presence of human macrophage colony-stimulating factor (MCS-F, 250 units/ml, Genetics Institute, Cambridge, MA). Differentiated macrophages were prepared for immunoprecipitation after 5 days in culture.

Bone marrow derived osteoclast-like cells were prepared as described by Wesolowski, *et al.*, 1995 *Exp. Cell Res.* 219:679-686, which is hereby incorporated by reference. After collagenase-dispase treatment, mononucleated tartrate resistant phosphatase positive cells were released from the tissue culture plate using 30 nM echistatin (Merck Res. Labs., West Point, PA). Freshly isolated osteoclast-like cells were immediately solubilized in immunoprecipitation buffer.

EXAMPLE 2

cDNA Cloning and Expression of mouse PYK2

Specific probes for mouse PYK2 and FAK were initially generated based on the non-homologous region between the proteins, which is adjacent to the C-terminal of the kinase domain. Using polymerase chain reaction (PCR), a specific probe for PYK2 (570bp) was generated using the 5'-primer AGTGA CATTT ATCAG ATGGA G (SEQ. ID. NO:1) and the 3'-primer GAATG GACTG TGCAC CGAGC C (SEQ. ID. NO:2) with cDNAs of mouse bone marrow derived osteoclast-like cells as template (Wesolowski, *et al.*, 1995, *supra*).

Similarly, a specific probe for FAK (700bp) was generated using the following primers: 5'- CAGCA CACAA TCCTG GAGGA G

(SEQ. ID. NO:3) and 3'- GCTGA AGCTT GACAC CCTCA T (SEQ. ID. NO. 4) with cDNAs of mouse osteoblastic MB1.8 cells as template (Wesolowski, *et al.*, 1995, *supra).* These probes were confirmed by sequencing analysis. PYK2 cDNA fragments were cloned from a mouse spleen λ-ZAP II cDNA library (Stratagene, La Jolla, CA) using the specific PYK2 probe. Full length PYK2 cDNA were constructed by ligation of two overlapping clones at the VspI site. The amino acid sequence of the isolated PYK2 cDNA clone was identical to the previously published mouse RAFTK sequence (Avraham, *et al.*, 1995 *supra*). Full length FAK cDNA was generated by PCR according to the published sequence as described in Hanks, *et al.*, 1992 *Proc. Natl. Acad. Sci. USA.* 89:8487-8491.

Both PYK2 and FAK cDNAs were subcloned into pCDNA3 plasmid (InVitrogen, San Diego, CA) and transfected into human embryonic kidney (HEK) 293 cells (ATCC, Rockland, MD) by electroporation at 200V, 960 μF using a GenePulser (Biorad Labs, Richmond, CA). HEK 293 cells were subsequently subjected to G418 selection (800 μg/ml, Gibco BRL) and clones were picked after 3 weeks in selection medium.

Expression of PYK2 and FAK in HEK293 cells were confirmed by Northern analysis using the respective probes, and Western blots were performed using anti-PYK2 antibodies. Mouse multiple tissue Northern blot was purchased from Clonetech (Palo Alto, CA) and hybridization of the Northern blot using probes specific for PYK2, FAK and glyceraldehde 3-phosphate dehydrogenase (GAPDH) were performed as described previously (Wesolowski, *et al.*, 1995, *supra*).

EXAMPLE 3

Production and Affinity Purification of Polyclonal Antibodies to mouse PYK2

The PYK2 C-terminal domain (from Methionine residue 685 to end) was amplified by PCR using the mouse PYK2 as template. Amplified product was cloned into plasmid pGEX-4T (Pharmacia Biotech., Piscataway, NJ) and transformed in *E.coli* XL1-Blue (Stratagene). Expression of GST-PYK2 C-terminal fragment was induced using 0.5 mM IPTG, purified and cleaved from GST with thrombin, essentially according to the instructions of the manufacturer (Pharmacia). The purified C-terminal fragment of mouse PYK2 was used to immunize two rabbits (Research Genetics, Huntsville, AL) and the titers of both antisera were initially determined by ELISA using the recombinant C-terminal fragment of PYK2. Specificity of the immune sera was subsequently determined by Western blot by comparison to the preimmune sera. Polyclonal antibodies were then affinity purified by passing the combined fractions of both antisera through an affinity column, which was constructed using the same purified antigen cross linked to CNBr-activated Sepharose 4B according to the instructions of the manufacturer (Pharmacia).

The antibodies were eluted from the column using 0.2 M Glycine, pH 2.5 and 1mM EGTA and the eluted fraction was then dialyzed against PBS containning 0.02% azide. Anti-PYK2 antibodies were stored at -70°C at a concentration of 0.5mg/ml.

EXAMPLE 4

In vitro Kinase Assay

After cell attachment to ECM, IC-21 cells were solubilized in TNE lysis buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1% NP-40, 1mM EDTA, 10% glycerol, 50 mM NaF, 1 mM sodium vanadate and protease inhibitors as described above. PYK2 were immunoprecipitated from the clarified lysates, half of the sample was subjected to immunoblotting with anti PYK2 antibodies, as described

above, and the other half was washed 2 times with the same lysis buffer, and with kinase assay buffer (1X) containing 20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 10 mM $MnCl_2$ and 1 mM dithiothreitol. After removal of the wash buffer, 50 µl of kinase assay buffer containing 5 µCi [γ-^{32}P] ATP (3000 Ci/mmol, Amersham), 10 mM ATP, 0.1% BSA and 100 µg of poly (Glu,Tyr) (molar ratio 4:1; Sigma) was added to the beads and incubated for 10 min at 30°C (Howell and Cooper, 1995 *Mol. Cell. Biol.* 14:5402-5411). The reaction mixtures (25 µl) were added to 25 µl of 30% trichloroacetic acid (TCA) and 0.1 M sodium pyrophosphate, followed by incubation at 4°C for 15 min. The precipitated proteins were transferred to a Multiscreen-FC filter plate (Millipore, Marlborough, MA), washed with ice cold 15% TCA (3X), allowed to dry and incorporation of ^{32}P into the substrate was counted on a Packard top count microplate scintillation counter (Packard, Meriden, CT). Each assay were performed as triplicate. The specific activity was determined by comparing the radioactive counts with immunoblot signals.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: DUONG, LE T.
RODAN, GIDEON A.

(ii) TITLE OF THE INVENTION: PROTEIN TYROSINE KINASE 2 (PYK2), NUCLEIC ACIDS AND ASSAY

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Merck & Co., Inc.
(B) STREET: P.O. Box 2000, 126 E. Lincoln Avenue
(C) CITY: Rahway
(D) STATE: NJ
(E) COUNTRY: USA
(F) ZIP: 07065-0900

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 60/037,561
(B) FILING DATE: 11-FEB-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sabatelli, Anthony D
(B) REGISTRATION NUMBER: 34,714
(C) REFERENCE/DOCKET NUMBER: 19792

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 732-594-1935
(B) TELEFAX: 732-594-4720
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGTGACATTT ATCAGATGGA G 21

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAATGGACTG TGCACCGAGC C 21

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAGCACACAA TCCTGGAGGA G 21

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCTGAAGCTT GACACCCTCA T 21

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3981 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATTCGCGGCC GCTCGACCTC AGCCTCTGCA GGCAGAGCCG CGCGTCCTAC CTGCGGCGGC 60
TGCGCTCACC TGGCCCAGCC CGGAGCCCTG GCCCGAGTCC GCGCCTCGCC CGAGGGACTG 120
CAATGTGCCG GTCCTAGCTG CAGTCTGAGA GGATGTCCGG GGTGTCTGAG CCCTTGAGCC 180
GTGTAAAAGT GGGCACTTTA CGCCGGCCTG AGGGCCCCCC AGAGCCCATG GTGGTGGTAC 240
CAGTGGATGT GGAGAAGGAA GACGTGCGCA TCCTCAAGGT CTGCTTCTAC AGCAACAGCT 300
TCAACCCAGG GAAGAACTTC AAGCTTGTCA AATGCACAGT GCAGACAGAG ATCCAGGAGA 360
TCATCACCTC CATCCTCCTG AGTGGGCGAA TAGGGCCCAA CATCCAGCTG GCTGAATGCT 420
ATGGGCTGAG GCTGAAGCAC ATGAAGTCAG ACGAGATCCA CTGGCTGCAC CCACAGATGA 480
CCGTGGGCGA AGTGCAGGAC AAGTATGAAT GTCTACACGT GGAAGCTGAG TGGAGGTATG 540
ACCTTCAAAT CCGCTACTTG CCGGAAGACT TCATGGAGAG CCTGAAAGAA GACAGGACCA 600
CATTGCTGTA CTTTTATCAA CAGCTCCGGA ATGACTACAT GCAACGCTAC GCCAGCAAGG 660
TCAGTGAAGG CATGGCTCTG CAGCTGGGCT GTCTGGAGCT CAGGAGATTC TTCAAGGACA 720
TGCCCCACAA TGCACTGGAC AAAAAGTCCA ACTTTGAACT CCTGGAAAAA GAAGTCGGTC 780
TGGACCTGTT TTTCCCAAAG CAGATGCAGG AAAACTTAAA GCCCAAGCAG TTCCGGAAGA 840
TGATCCAGCA GACCTTCCAG CAGTATGCAT CACTCCGGGA GGAAGAGTGT GTCATGAAAT 900
TCTTCAATAC CCTAGCGGGC TTTGCCAACA TTGACCAGGA GACCTACCGC TGCGAACTCA 960
TTCAAGGATG GAACATTACT GTGGACCTGG TCATCGGCCC TAAAGGCATC CGTCAGCTGA 1020
CAAGTCAAGA TACAAAGCCC ACCTGCCTGG CCGAGTTTAA GCAGATCAGA TCCATCAGGT 1080
GCCTCCCATT GGAAGAGACC CAGGCAGTCC TGCAGCTGGG CATCGAGGGT GCCCCCCAGT 1140
CCTTGTCTAT CAAAACGTCG TCCCTGGCAG AGGCTGAGAA CATGGCTGAT CTCATAGATG 1200
GCTACTGCAG GCTGCAAGGA GAACATAAGG GCTCTCTCAT CATGCATGCC AAGAAAGATG 1260
GTGAGAAGAG GAACAGCCTG CCTCAGATCC CCACACTAAA CCTGGAGGCT CGGCGGTCGC 1320
ACCTCTCAGA AAGCTGCAGC ATAGAGTCAG ACATCTATGC GGAGATTCCC GATGAGACCC 1380
TGCGAAGACC AGGAGGTCCA CAGTACGGTG TTGCCCGTGA AGAAGTAGTT CTTAACCGCA 1440
TTCTGGGTGA AGGCTTCTTT GGGGAGGTCT ATGAAGGTGT CTACACGAAC CACAAAGGGG 1500
AAAAAATTAA TGTGGCCGTC AAGACCTGTA AGAAAGACTG TACCCAGGAC AACAAGGAGA 1560
AGTTCATGAG TGAGGCAGTG ATCATGAAGA ATCTTGACCA CCCTCACATC GTGAAGCTGA 1620
TTGGCATCAT TGAAGAGGAA CCCACCTGGA TTATCATGGA ACTGTATCCT TATGGGGAGC 1680
TGGGACACTA CCTGGAACGA AATAAAAACT CCCTGAAGGT ACCCACTCTG GTCCTGTACA 1740
CCCTACAGAT ATGCAAAGCC ATGGCCTATC TGGAGAGCAT CAACTGTGTG CACAGGGATA 1800
TTGCTGTCCG GAACATCCTG GTGGCCTCTC CTGAGTGTGT GAAGCTGGGG GACTTTGGGC 1860
TCTCCCGGTA CATTGAGGAC GAAGACTATT ACAAAGCCTC TGTGACACGT CTACCCATCA 1920
AATGGATGTC CCCCGAGTCC ATCAACTTCC GCCGCTTCAC AACCGCCAGT GATGTCTGGA 1980
TGTTTGCTGT ATGCATGTGG GAGATCCTCA GCTTTGGGAA GCAGCCTTTC TTCTGGCTCG 2040
AAAATAAGGA TGTCATCGGA GTGCTGGAGA AAGGGGACAG GCTGCCCAAG CCCGAACTCT 2100
GTCCGCCTGT CCTTTACACA CTCATGACTC GCTGCTGGGA CTACGACCCC AGTGACCGGC 2160
CCCGCTTCAC GGAGCTTGTG TGCAGCCTCA GTGACATTTA TCAGATGGAG AAGGACATTG 2220
CCATAGAGCA AGAAAGGAAT GCTCGCTACC GACCCCCTAA AATATTGGAG CCTACTACCT 2280
TTCAGGAACC CCCACCCAAG CCCAGCCGGC CCAAGTACAG ACCTCCTCCA CAGACCAACC 2340
TGCTGGCTCC TAAGCTGCAG TTCCAGGTCC CTGAGGGTCT GTGTGCCAGC TCTCCTACGC 2400
TTACCAGCCC TATGGAGTAT CCATCTCCAG TTAACTCGCT GCACACCCCA CCTCTCCACC 2460
GGCACAATGT CTTCAAGCGC CACAGCATGC GGGAGGAGGA CTTCATCCGG CCCAGTAGCC 2520
GAGAAGAGGC CCAGCAGCTC TGGGAGGCAG AGAAGATCAA GATGAAGCAG GTCCTAGAAA 2580
GACAGCAGAA GCAGATGGTG GAAGATTCCC AGTGGCTGAG GCGAGAGGAA AGATGCTTGG 2640
ACCCTATGGT TTATATGAAT GACAAGTCCC CACTGACTCC AGAGAAGGAG GCCGGCTACA 2700
CGGAGTTCAC AGGGCCCCCA CAGAAACCAC CTCGGCTCGG TGCACAGTCC ATTCAGCCCA 2760
CAGCCAACCT GGACAGGACC GATGACCTCG TGTACCACAA TGTCATGACC CTGGTGGAGG 2820
CTGTGCTGGA ACTCAAGAAC AAGCTTGGCC AGTTGCCCCC TGAGGACTAT GTGGTGGTGG 2880
TGAAGAACGT GGGGCTGAAC CTGCGGAAGC TCATCGGCAG TGTGGACGAT CTCTTGCCCT 2940
CCTTGCCGGC ATCTTCGAGG ACAGAGATTG AAGGGACCCA GAAACTGCTC AACAAAGACC 3000
TGGCAGAGCT CATCAACAAG ATGAAGTTGG CTCAGCAGAA CGCCGTGACG TCCCTGAGTG 3060
AGGACTGCAA GCGGCAGATG CTCACAGCGT CCCATACCCT GGCTGTGGAT GCCAAGAACC 3120
TGCTGGATGC TGTGGACCAA GCCAAGGTTG TGGCTAATCT GGCCCACCCG CCTGCAGAGT 3180
GATCAAGAGA GGGGCCACCT GCCTGCATCT TCTGCCCCCA CCTGTCTTGG CATACCTTTC 3240

```
CTGCCTTGCC TTTGGTTATT GGTCTTCCAG GGAAAGCTGA GAAGAGTCCA TCCCCCTTGC   3300
CACTTTGCAC GACACCCCCT CTTCCCCCAA CCCACCCCAG ACTGTGCTAC TCAGGCTGCA   3360
TCTGGACAGA AAGGACTCTG GGCACAGACA CGGGGTGGGG TGACATAGTT CATAGGGGTA   3420
CTACTGCCAG CCACTCCCTC TTACCCCAGC CTGGGTTGCT GGAGCATCAT TGGGGTCATG   3480
AGTGTACCCC TAACGGCCAA GATGGCTTTC TGCATGGACA TTTGAGAGCC AGTATTCCTC   3540
CTTCCTCTTC AGCCCTCAGG GACCCCTGAT ACAGAGGGGA CAGAGAGGGG TTTTATTTGT   3600
AGAGAAGCTG GTGAGATGAG GGCTGGACCT GGCTCTCTTG TACAGTGTAC ATTGGAATTT   3660
ATTTAATGTG AGTTTGACCT GGATGGACAG CCAAGGGCCA TAGTCCAGGA GCAAACCAAT   3720
CCAGTCACAG GACTCTGTGT TTTATGGAAC TGAGTGCCAC AGGAAGAAGC AGAGAGTCGG   3780
AGGTCAGAAT GGACTTTGTG CCCTTCCTGC GTTTCTCTTC TCCCTCTTTC CTTCTCCCCT   3840
CTTTTCTTAC GTCTCCTTTT TCTCCTCCCC CTTTTCACAT CTGCTCCCCT CCTCTCTCAT   3900
GTCTGTGGAG AACATTTACC TTCCTTCTTT TTGATCGGTG GTTGAATTAA AATTATTACC   3960
ATTTGCTTTG TGGCTCGTGC C                                             3981
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1009 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Val Gly Thr Leu
 1               5                  10                  15
Arg Arg Pro Glu Gly Pro Pro Glu Pro Met Val Val Val Pro Val Asp
            20                  25                  30
Val Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn
        35                  40                  45
Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln
    50                  55                  60
Thr Glu Ile Gln Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile
65                  70                  75                  80
Gly Pro Asn Ile Gln Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His
                85                  90                  95
Met Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly
            100                 105                 110
Glu Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg
        115                 120                 125
Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu
    130                 135                 140
Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn
145                 150                 155                 160
Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val Ser Glu Gly Met Ala Leu
                165                 170                 175
Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe Phe Lys Asp Met Pro His
            180                 185                 190
Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu Leu Leu Glu Lys Glu Val
        195                 200                 205
Gly Leu Asp Leu Phe Phe Pro Lys Gln Met Gln Glu Asn Leu Lys Pro
    210                 215                 220
Lys Gln Phe Arg Lys Met Ile Gln Gln Thr Phe Gln Gln Tyr Ala Ser
225                 230                 235                 240
```

```
Leu Arg Glu Glu Glu Cys Val Met Lys Phe Phe Asn Thr Leu Ala Gly
                245                 250                 255
Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg Cys Glu Leu Ile Gln Gly
            260                 265                 270
Trp Asn Ile Thr Val Asp Leu Val Ile Gly Pro Lys Gly Ile Arg Gln
        275                 280                 285
Leu Thr Ser Gln Asp Thr Lys Pro Thr Cys Leu Ala Glu Phe Lys Gln
    290                 295                 300
Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu Glu Thr Gln Ala Val Leu
305                 310                 315                 320
Gln Leu Gly Ile Glu Gly Ala Pro Gln Ser Leu Ser Ile Lys Thr Ser
                325                 330                 335
Ser Leu Ala Glu Ala Glu Asn Met Ala Asp Leu Ile Asp Gly Tyr Cys
            340                 345                 350
Arg Leu Gln Gly Glu His Lys Gly Ser Leu Ile Met His Ala Lys Lys
        355                 360                 365
Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln Ile Pro Thr Leu Asn Leu
    370                 375                 380
Glu Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp
385                 390                 395                 400
Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro
                405                 410                 415
Gln Tyr Gly Val Ala Arg Glu Glu Val Val Leu Asn Arg Ile Leu Gly
            420                 425                 430
Glu Gly Phe Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys
        435                 440                 445
Gly Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr
    450                 455                 460
Gln Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn
465                 470                 475                 480
Leu Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu
                485                 490                 495
Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His
            500                 505                 510
Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Pro Thr Leu Val Leu
        515                 520                 525
Tyr Thr Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn
    530                 535                 540
Cys Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro
545                 550                 555                 560
Glu Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp
                565                 570                 575
Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met
            580                 585                 590
Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val
        595                 600                 605
Trp Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln
    610                 615                 620
Pro Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys
625                 630                 635                 640
Gly Asp Arg Leu Pro Lys Pro Glu Leu Cys Pro Pro Val Leu Tyr Thr
                645                 650                 655
Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe
            660                 665                 670
Thr Glu Leu Val Cys Ser Leu Ser Asp Ile Tyr Gln Met Glu Lys Asp
        675                 680                 685
```

```
Ile Ala Ile Glu Gln Glu Arg Asn Ala Arg Tyr Arg Pro Pro Lys Ile
    690                 695                 700
Leu Glu Pro Thr Thr Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro
705                 710                 715                 720
Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln
                725                 730                 735
Phe Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser
            740                 745                 750
Pro Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu
        755                 760                 765
His Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe
    770                 775                 780
Ile Arg Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu
785                 790                 795                 800
Lys Ile Lys Met Lys Gln Val Leu Glu Arg Gln Gln Lys Gln Met Val
                805                 810                 815
Glu Asp Ser Gln Trp Leu Arg Arg Glu Glu Arg Cys Leu Asp Pro Met
            820                 825                 830
Val Tyr Met Asn Asp Lys Ser Pro Leu Thr Pro Glu Lys Glu Ala Gly
        835                 840                 845
Tyr Thr Glu Phe Thr Gly Pro Pro Gln Lys Pro Pro Arg Leu Gly Ala
    850                 855                 860
Gln Ser Ile Gln Pro Thr Ala Asn Leu Asp Arg Thr Asp Asp Leu Val
865                 870                 875                 880
Tyr His Asn Val Met Thr Leu Val Glu Ala Val Leu Glu Leu Lys Asn
                885                 890                 895
Lys Leu Gly Gln Leu Pro Pro Glu Asp Tyr Val Val Val Val Lys Asn
            900                 905                 910
Val Gly Leu Asn Leu Arg Lys Leu Ile Gly Ser Val Asp Asp Leu Leu
        915                 920                 925
Pro Ser Leu Pro Ala Ser Ser Arg Thr Glu Ile Glu Gly Thr Gln Lys
    930                 935                 940
Leu Leu Asn Lys Asp Leu Ala Glu Leu Ile Asn Lys Met Lys Leu Ala
945                 950                 955                 960
Gln Gln Asn Ala Val Thr Ser Leu Ser Glu Asp Cys Lys Arg Gln Met
                965                 970                 975
Leu Thr Ala Ser His Thr Leu Ala Val Asp Ala Lys Asn Leu Leu Asp
            980                 985                 990
Ala Val Asp Gln Ala Lys Val Val Ala Asn Leu Ala His Pro Pro Ala
        995                1000                1005
Glu
  1
```

WHAT IS CLAIMED IS:

1. A nucleic acid, free from associated nucleic acids which encodes murine protein tyrosine kinase 2 (PYK2).

2. A nucleic acid according to Claim 1 which is DNA.

3. Murine PYK2 cDNA.

4. Murine PYK2 cDNA which is set forth in Figure 1.

5. A cell line comprising heterologous PYK2, and which expresses PYK2.

6. An assay to identify compounds which alter the activity of PYK2 comprising:

a) contacting recombinant PYK2 with a tyrosine substrate in the presence of radiolabeled ATP and a putative activity-modifying compound;

b) measuring the amount of radiolabeled tyrosine which is formed; and

c) optionally comparing the amount of radiolabeled tyrosine formed in the presence of the putative activity-modifying compound with that formed in the absence of the putative activity-modifying compound.

TITLE OF THE INVENTION

PROTEIN TYROSINE KINASE 2 (PYK2), NUCLEIC ACIDS, AND ASSAY

ABSTRACT OF THE INVENTION

This invention is directed to nucleic acids encoding protein tyrosine kinase 2 (PYK2), to murine PYK2, to methods of making this protein using the nucleic acids, and to assays for inhibitors of PYK2.

```
attcgcggccgctcgacctcagcctctgcaggcagagccgcgcgtcctacctgcggcggc    60
tgcgctcacctggcccagcccggagccctggcccgagtccgcgcctcgcccgagggactg   120
caatgtgccggtcctagctgcagtctgagaggATGTCCGGGGTGTCTGAGCCCTTGAGCC   180
                                    M  S  G  V  S  E  P  L  S  R    10
GTGTAAAAGTGGGCACTTTACGCCGGCCTGAGGGCCCCCCAGAGCCCATGGTGGTGGTAC   240
  V  K  V  G  T  L  R  R  P  E  G  P  P  E  P  M  V  V  V  P    30
CAGTGGATGTGGAGAAGGAAGACGTGCGCATCCTCAAGGTCTGCTTCTACAGCAACAGCT   300
  V  D  V  E  K  E  D  V  R  I  L  K  V  C  F  Y  S  N  S  F    50
TCAACCCAGGGAAGAACTTCAAGCTTGTCAAATGCACAGTGCAGACAGAGATCCAGGAGA   360
  N  P  G  K  N  F  K  L  V  K  C  T  V  Q  T  E  I  Q  E  I    70
TCATCACCTCCATCCTCCTGAGTGGGCGAATAGGGCCCAACATCCAGCTGGCTGAATGCT   420
  I  T  S  I  L  L  S  G  R  I  G  P  N  I  Q  L  A  E  C  Y    90
ATGGGCTGAGGCTGAAGCACATGAAGTCAGACGAGATCCACTGGCTGCACCCACAGATGA   480
  G  L  R  L  K  H  M  K  S  D  E  I  H  W  L  H  P  Q  M  T   110
CCGTGGGCGAAGTGCAGGACAAGTATGAATGTCTACACGTGGAAGCTGAGTGGAGGTATG   540
  V  G  E  V  Q  D  K  Y  E  C  L  H  V  E  A  E  W  R  Y  D   130
ACCTTCAAATCCGCTACTTGCCGGAAGACTTCATGGAGAGCCTGAAAGAAGACAGGACCA   600
  L  Q  I  R  Y  L  P  E  D  F  M  E  S  L  K  E  D  R  T  T   150
CATTGCTGTACTTTTATCAACAGCTCCGGAATGACTACATGCAACGCTACGCCAGCAAGG   660
  L  L  Y  F  Y  Q  Q  L  R  N  D  Y  M  Q  R  Y  A  S  K  V   170
TCAGTGAAGGCATGGCTCTGCAGCTGGGCTGTCTGGAGCTCAGGAGATTCTTCAAGGACA   720
  S  E  G  M  A  L  Q  L  G  C  L  E  L  R  R  F  F  K  D  M   190
```

FIG. 1A

```
TGCCCCACAATGCACTGGACAAAAAGTCCAACTTTGAACTCCTGGAAAAAGAAGTCGGTC   780
  P  H  N  A  L  D  K  K  S  N  F  E  L  L  E  K  E  V  G  L   210
TGGACCTGTTTTTCCCAAAGCAGATGCAGGAAAACTTAAAGCCCAAGCAGTTCCGGAAGA   840
  D  L  F  F  P  K  Q  M  Q  E  N  L  K  P  K  Q  F  R  K  M   230
TGATCCAGCAGACCTTCCAGCAGTATGCATCACTCCGGGAGGAAGAGTGTGTCATGAAAT   900
  I  Q  Q  T  F  Q  Q  Y  A  S  L  R  E  E  E  C  V  M  K  F   250
TCTTCAATACCCTAGCGGGCTTTGCCAACATTGACCAGGAGACCTACCGCTGCGAACTCA   960
  F  N  T  L  A  G  F  A  N  I  D  Q  E  T  Y  R  C  E  L  I   270
TTCAAGGATGGAACATTACTGTGGACCTGGTCATCGGCCCTAAAGGCATCCGTCAGCTGA  1020
  Q  G  W  N  I  T  V  D  L  V  I  G  P  K  G  I  R  Q  L  T   290
CAAGTCAAGATACAAAGCCCACCTGCCTGGCCGAGTTTAAGCAGATCAGATCCATCAGGT  1080
  S  Q  D  T  K  P  T  C  L  A  E  F  K  Q  I  R  S  I  R  C   310
GCCTCCCATTGGAAGAGACCCAGGCAGTCCTGCAGCTGGGCATCGAGGGTGCCCCCCAGT  1140
  L  P  L  E  E  T  Q  A  V  L  Q  L  G  I  E  G  A  P  Q  S   330
CCTTGTCTATCAAAACGTCGTCCCTGGCAGAGGCTGAGAACATGGCTGATCTCATAGATG  1200
  L  S  I  K  T  S  S  L  A  E  A  E  N  M  A  D  L  I  D  G   350
GCTACTGCAGGCTGCAAGGAGAACATAAGGGCTCTCTCATCATGCATGCCAAGAAAGATG  1260
  Y  C  R  L  Q  G  E  H  K  G  S  L  I  M  H  A  K  K  D  G   370
GTGAGAAGAGGAACAGCCTGCCTCAGATCCCCACACTAAACCTGGAGGCTCGGCGGTCGC  1320
  E  K  R  N  S  L  P  Q  I  P  T  L  N  L  E  A  R  R  S  H   390
ACCTCTCAGAAAGCTGCAGCATAGAGTCAGACATCTATGCGGAGATTCCCGATGAGACCC  1380
  L  S  E  S  C  S  I  E  S  D  I  Y  A  E  I  P  D  E  T  L   410
```

FIG. 1B

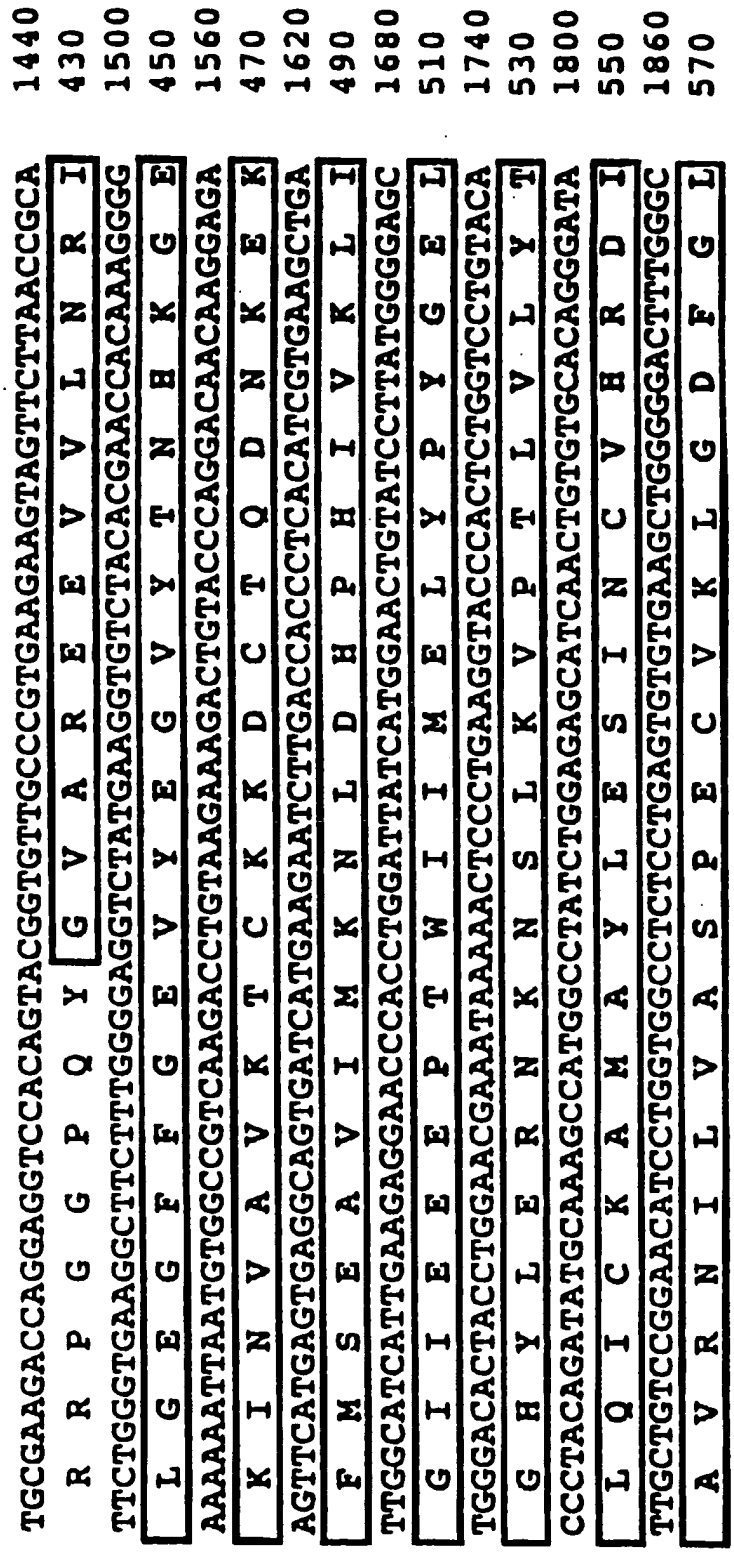
TGCGAAGACCAGGAGGTCCACAGTACGGTGTTGCCCGTGAAGAAGTAGTTCTTAACCGCA 1440
R R P G G G P Q Y G V A R E E E V V L N R I 430
TTCTGGGTGAAGGCTTCTTTGGGGAGGTCTATGAAGGTGTCTACACGAACCACAAAGGGG 1500
L G E G F F G E V Y E G V Y T N H K G E 450
AAAAATTAATGTGGCCGTCAAGACCTGTAAGAAAGACTGTACCCAGGACAACAAGGAGA 1560
K I N V A V K T C K K D C T Q D N K E K 470
AGTTCATGAGTGAGGCAGTGATCATGAAGAATCTTGACCACCCTCACATCGTGAAGCTGA 1620
F M S E A V I M K N L D H P H I V K L I 490
TTGGCATCATTGAGAGGAACCCACCTGGATTATCATGGAACTGTATCCTTATGGGGAGC 1680
G I I E E P T W I I M E L Y P Y G E L 510
TGGGACACTACCTGGAACGAAATAAAACTCCCTGAAGGTACCCACTCTGGTCCTGTACA 1740
G H Y L E R N K N S L K V P T L V L Y T 530
CCCTACAGATATGCAAAGCCATGGCCTATCTGGAGAGCATCAACTGTGTGCACAGGGATA 1800
L Q I C K A M A Y L E S I N C V H R D I 550
TTGCTGTCCGGAACATCCTGGTGGCCTCTCCTGAGTGTGAAGCTGGGGGACTTTGGGC 1860
A V R N I L V A S P E C V K L G D F G L 570

FIG. 1C

```
TCTCCCGGTACATTGAGGACGAAGACTATTACAAAGCCTCTGTGACACGTCTACCCATCA  1920
 S  R  Y  I  E  D  E  D  Y  Y  K  A  S  V  T  R  L  P  I  K   590
AATGGATGTCCCCCGAGTCCATCAACTTCCGCCGCTTCACAACCGCCAGTGATGTCTGGA  1980
 W  M  S  P  E  S  I  N  F  R  R  F  T  T  A  S  D  V  W  M   610
TGTTTGCTGTATGCATGTGGGAGATCCTCAGCTTTGGGAAGCAGCCTTTCTTCTGGCTCG  2040
 F- A  V  C  M  W  E  I  L  S  F  G  K  Q  P  F  F  W  L  E   630
AAAATAAGGATGTCATCGGAGTGCTGGAGAAAGGGGACAGGCTGCCCAAGCCCGAACTCT  2100
 N  K  D  V  I  G  V  L  E  K  G  D  R  L  P  K  P  E  L  C   650
GTCCGCCTGTCCTTTACACACTCATGACTCGCTGCTGGGACTACGACCCCAGTGACCGGC  2160
 P  P  V  L  Y  T  L  M  T  R  C  W  D  Y  D  P  S  D  R  P   670
CCCGCTTCACGGAGCTTGTGTGCAGCCTCAGTGACATTTATCAGATGGAGAAGGACATTG  2220
 R  F  T  E  L  V  C  S  L  S  D  I  Y  Q  M  E  K  D  I  A   690
CCATAGAGCAAGAAAGGAATGCTCGCTACCGACCCCCTAAAATATTGGAGCCTACTACCT  2280
 I  E  Q  E  R  N  A  R  Y  R  P  P  K  I  L  E  P  T  T  F   710
TTCAGGAACCCCCACCCAAGCCCAGCCGGCCCAAGTACAGACCTCCTCCACAGACCAACC  2340
 Q  E  P  P  P  K  P  S  R  P  K  Y  R  P  P  P  Q  T  N  L   730
TGCTGGCTCCTAAGCTGCAGTTCCAGGTCCCTGAGGGTCTGTGTGCCAGCTCTCCTACGC  2400
 L  A  P  K  L  Q  F  Q  V  P  E  G  L  C  A  S  S  P  T  L   750
TTACCAGCCCTATGGAGTATCCATCTCCAGTTAACTCGCTGCACACCCCACCTCTCCACC  2460
 T  S  P  M  E  Y  P  S  P  V  N  S  L  H  T  P  P  L  H  R   770
GGCACAATGTCTTCAAGCGCCACAGCATGCGGGAGGAGGACTTCATCCGGCCCAGTAGCC  2520
 H  N  V  F  K  R  H  S  M  R  E  E  D  F  I  R  P  S  S  R   790
```

FIG. 1D

```
GAGAAGAGGCCCAGCAGCTCTGGGAGGCAGAGAAGATCAAGATGAAGCAGGTCCTAGAAA  2580
  E  E  A  Q  Q  L  W  E  A  E  K  I  K  M  K  Q  V  L  E  R    810
GACAGCAGAAGCAGATGGTGGAAGATTCCCAGTGGCTGAGGCGAGAGGAAAGATGCTTGG  2640
  Q  Q  K  Q  M  V  E  D  S  Q  W  L  R  R  E  E  R  C  L  D    830
ACCCTATGGTTTATATGAATGACAAGTCCCCACTGACTCCAGAGAAGGAGGCCGGCTACA  2700
  P  M  V  Y  M  N  D  K  S  P  L  T  P  E  K  E  A  G  Y  T    850
CGGAGTTCACAGGGCCCCCACAGAAACCACCTCGGCTCGGTGCACAGTCCATTCAGCCCA  2760
  E  F  T  G  P  P  Q  K  P  P  R  L  G  A  Q  S  I  Q  P  T    870
CAGCCAACCTGGACAGGACCGATGACCTCGTGTACCACAATGTCATGACCCTGGTGGAGG  2820
  A  N  L  D  R  T  D  D  L  V  Y  H  N  V  M  T  L  V  E  A    890
CTGTGCTGGAACTCAAGAACAAGCTTGGCCAGTTGCCCCCTGAGGACTATGTGGTGGTGG  2880
  V  L  E  L  K  N  K  L  G  Q  L  P  P  E  D  Y  V  V  V  V    910
TGAAGAACGTGGGGCTGAACCTGCGGAAGCTCATCGGCAGTGTGGACGATCTCTTGCCCT  2940
  K  N  V  G  L  N  L  R  K  L  I  G  S  V  D  D  L  L  P  S    930
CCTTGCCGGCATCTTCGAGGACAGAGATTGAAGGGACCCAGAAACTGCTCAACAAAGACC  3000
  L  P  A  S  S  R  T  E  I  E  G  T  Q  K  L  L  N  K  D  L    950
TGGCAGAGCTCATCAACAAGATGAAGTTGGCTCAGCAGAACGCCGTGACGTCCCTGAGTG  3060
  A  E  L  I  N  K  M  K  L  A  Q  Q  N  A  V  T  S  L  S  E    970
AGGACTGCAAGCGGCAGATGCTCACAGCGTCCCATACCCTGGCTGTGGATGCCAAGAACC  3120
  D  C  K  R  Q  M  L  T  A  S  H  T  L  A  V  D  A  K  N  L    990
TGCTGGATGCTGTGGACCAAGCCAAGGTTGTGGCTAATCTGGCCCACCCGCCTGCAGAGt  3180
  L  D  A  V  D  Q  A  K  V  V  A  N  L  A  H  P  P  A  E  *
```

FIG. 1E

```
gatcaagagaggggccacctgcctgcatcttctgccccacctgtcttggcataccttc  3240
ctgccttgcctttggttattggtcttccagggaaagctgagaagagtccatccccttgc  3300
cactttgcacgacacccccttcccccaacccccagactgtgctactcaggctgca      3360
tctggacagaaaggactctgggcacagacacgggggtgggtgacatagttcataggggta 3420
ctactgccagccactccctcttaccccagcctgggttgctggagcatcattggggtcatg  3480
agtgtaccccctaacggccaagatggctttctgcatggacatttgagagccagtattcctc 3540
cttcctcttcagccccctcagggaccccctgatacagaggggacagaggagggtttttatttgt 3600
agagaagctggtgagatgagggctggacctggctcttcttgtacagtgtacattggaattt 3660
atttaatgtgagtttgacctggatggacagccaagggccatagtccaggagcaaaccaat  3720
ccagtcacaggactctgtgtttatggaactgagtgccacaggaagaagcagagagtcgg  3780
aggtcagaatggacttttgtgccctttcctgcgtttctcttccctctttcctcccct    3840
ctttcttacgtctcctttttctcctcccctttcacatctgctccccctcctctcat     3900
gtctgtggagaacatttacctttcctttttgatcggttgaattaaaattattacc      3960
atttgctttgtggctcgtgcc                                        3981
```

FIG. 1F

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/02494

A. CLASSIFICATION OF SUBJECT MATTER
IPC(6) :C12N 1/00, 5/10, 15/54; C12Q 1/48, 1/68
US CL : 536/23.2; 435/6, 252.3, 254.11, 325, 410
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.2; 435/4, 6, 7.4, 15, 194. 252.3, 254.11, 320.1, 325, 410

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	AVRAHAM et al. Identification and Characterization of a Novel Related Adhesion Focal Tyrosine Kinase (RAFTK) from Megakaryocytes and Brain. Journal of Biological Chemistry. 17 November 1995, Vol. 270, No. 46, pages 27742-27751, especially Figure 2, row 2.	1-4 5-6
Y	US 5,573,944 A (KIRSCHNER ET AL.) 12 November 1996, column 2, lines 52-63; column 7, lines 50-55; claims 20-23.	5
Y	US 5,538,858 A (MALIA ET AL.) 23 July 1996, claims 1-3, 7, 10	6

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 26 MARCH 1998	Date of mailing of the international search report 01 JUN 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer GABRIELE E. BUGAISKY Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/02494

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, STN-CAS FILES REGISTRY, MEDLINE, CAPLUS, N-GENESEQ, GENBANK/EMBL
search terms: protein tyrosine kinase, pyk2, pyk 2, mouse, murine, mus, raftk, cadtk, cakbeta, cell adhesion kinase beta, related adhesion focal tyrosine kinase, inhibit?, modulat?, antagoni?, agoni?